

基于神经网络和综合特征的车牌定位算法

王 森, 陈炬桦

(中山大学, 广东 广州 510275)

摘 要:文中提出一种基于神经网络,利用车牌颜色、字符分布特征来提取车牌的算法。与以前的神经网络定位车牌不同的是,本算法是用二值化后每个8-连接对象作为网络的输入。这样可以减少训练样本数目,有针对性地训练噪音。实验证明本算法对于复杂背景的车牌有较好的提取效果,并且有较快的执行速度和较好的鲁棒性。

关键词:车牌定位;神经网络;颜色空间;灰度共生矩阵

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2008)0038-04

Algorithm of Car Plate Location Based on Neural Network and Integrated Features

WANG Sen, CHEN Ju-hua

(Sun Yat-sen University, Guangzhou 510275, China)

Abstract: Presents an algorithm which makes use of information such as color of the car plate and the distribution of the car plate characters based on neural network to extract the car plate. The algorithm uses every 8-connection objects as the input of the neural network which is different from other neural network location of car plate. The method could reduce the sample used for training the network and training the noise of the picture directly. Experiments show that method is effective and robust for the photo which have complicated backgrounds and meantime execute very fast.

Key words: car plate location; neural network; color space; gray level co-occurrence matrix

0 前 言

车牌识别是智能交通系统(ITS)中最核心的技术,也是目前智能交通的热点研究课题。一般来说,车牌识别主要分为三个步骤:车牌定位、字符分割和字符识别。车牌定位是车牌识别系统中首要解决的也是最关键的部分,它为后续的字符分割以及识别部分奠定基础。国内外对车牌定位的研究很多,大致可以分为三种方法:

- * 利用边缘信息进行定位。
- * 利用车牌色彩信息进行定位^[1]。
- * 利用神经网络定位^[2,3]。

很多方法都具有较好的效果并得到了实际应用。

文中提出一种基于神经网络的车牌定位算法。神经网络具有抗噪、容错、自学能力强、并行性等诸多优点。而且经实验可知,只要充分地训练网络,就可以将

复杂的背景噪音以及车牌字符正确归类,达到车牌分割的目的。笔者的算法采用研究最成熟的BP神经网络来实现。与以前的神经网络定位不同的是:将二值化后的图像的每个8-连接对象经过归一化处理后作为神经网络的输入。这里首先要介绍一个定义:8-连接和4-连接^[4]。为此引入邻域的概念。

定义1 近邻:在二值图像中每个像素 P_i 的近邻定义为与 P_i 相邻的最近的8个像素。如表1所示,奇数下标的4个像素表示 P_i 的4-近邻,除 P_i 外与 P_i 相邻的所有8个像素称为 P_i 的8-近邻。

表1 像素 P_i 的近邻

K_4	K_3	K_2
K_5	P_i	K_1
K_6	K_7	K_8

定义2 连接:在二值图像中的一系列像素 $p_1 p_2 \cdots p_n$,若每个 p_i 均是 p_{i+1} ($1 \leq i \leq n-1$)的4-近邻(或8-近邻),则称 $p_1 p_2 \cdots p_n$ 组成的对象为4-连接(或8-连接)的。算法的步骤如下:

1) 图像预处理:对原图进行灰度化、灰度拉伸、形态开运算、图像二值化处理,得 P 。

收稿日期:2007-05-25

作者简介:王 森(1982-),男,河南汤阴人,硕士研究生,研究方向为并行与分布式计算;陈炬桦,博士后,副教授,研究方向为并行与分布式计算、计算机软件理论。

2)BP神经网络粗定位:对 P 标识每个8-连接的对象,神经网络输出为高信号的对象为可能的车牌字符。将这些对象记录到一个数组 C 内。

3)利用综合特征精确定位:对 C 内的每个对象 C_i 进行如下规则的分析:

(1)HSV色彩空间分析;

(2) C_i 一定范围内必定有三个以上 C 中的元素,这些元素记为数组 P ;

(3) P 中元素的高应该比较接近,其方差小于某个阈值 ε 。

1 BP神经网络

反向传播网络(Back-Propagation Network,简称BP网络)是人工神经网络最精华的部分,也是前向网络的核心部分。它要求激活函数必须是连续可微的。其激活函数所划分的区域是非线性的超平面组成的区域。所以理论上BP网络可以逼近任意非线性函数。

1.1 网络的结构

BP网络实质就是多层感知器(Multi-Layer Perceptron,MLP),对于常用的三层(含有输入层、隐含层和输出层)网络来说,第一层属于输入层,接受输入向量;第二层属于隐含层,用于记忆,增加网络的可调参数,使网络输出更加精确;第三层属于输出层,输出网络结果。相邻层之间的节点属于全连接,相同层之间的节点没有连接。理论上对于三层BP网络只要使得对隐含层的节点增大到一定的范围就可以拟合任何非线性函数。也就是说一个隐含层和多个隐含层的效果是相同的。网络常用的激活函数有两个:logsig和tansig,其表达式为:

$$\begin{aligned}\text{logsig}(n) &= \frac{1}{(1 + \exp(-n))} \\ \text{tansig}(n) &= \frac{2}{(1 + \exp(-2 * n))} - 1\end{aligned}\quad (1)$$

1.2 网络的学习

网络由于其学习算法采用BP算法所以又被称为BP网络,BP算法属于监督式的算法。对于 q 个输入学习样本: p^1, p^2, \dots, p^q ,已知与其对应的网络输出样本为: T^1, T^2, \dots, T^q ,学习的目的是使网络的实际输出 $A^1 A^2 \dots A^q$ 与目标矢量 T^1, T^2, \dots, T^q 之间的误差达到期望值。误差函数描述为:

$$E(W, B) = \frac{1}{2} \sum_{k=1}^n (t_k - a_{2k})^2 \quad (2)$$

其中 a_{2k} 为实际的输出, t_k 为期望输出。BP算法学习分两个阶段:第一阶段是前向计算阶段,这个阶段权值不会改变,输出矢量 $A^1 A^2 \dots A^q$ 与目标矢量 $T^1, T^2,$

\dots, T^q 输入存在一定的误差。第二个阶段是误差反向传播阶段,本阶段通过误差的层层传播来修改每层的权值,使得误差信号变小,最后通过多次迭代达到误差期望值。每层的权值和偏差可以看作误差函数的变量,BP学习的实质是通过改变权值和偏差,使得误差函数值沿其梯度方向下降^[5],称为梯度下降法。

2 车牌定位

2.1 图像预处理

对一幅RGB图像首先进行灰度化处理。公式如下:

$$R = G = B = \frac{W_R * R + W_G * G + W_b * B}{3} \quad (3)$$

其中, W_R, W_G, W_b 分别为 R, G, B 的权。 $W_R = 0.30, W_G = 0.59, W_b = 0.11$ 。对灰度化后的图像采用局部灰度拉伸,使得车牌字符的灰度值与背景的灰度值的差变大,有利于后续的图像二值化。Matlab中衡量拉伸效果的为一个参数值 r, r 大于1可以对高灰度值进行很好的拉伸,小于1可以对低灰度值进行很好的拉伸,等于1则是线性拉伸。采用不同的拉伸参数得出的二值化图像不同,车牌字符的连通性也不同。为了得到较理想的拉伸参数做了以下试验:

利用文中所说的算法,采用不同的拉伸参数(0.6~3.0)对100幅照片分别进行了25次的定位。有2幅照片未正确的定位,共给出14个错误的车牌区域。所有照片共得出1614个车牌结果,错误的有168个,去除未正确定位车牌照片的7个结果,也就是说可以定位车牌的照片中共有161个错误结果,车牌定位的置信度为90%,准确率为98%。拒识率为0%。也就是说只要采用正确的拉伸参数,就可以达到98%定位的准确率。

笔者实验发现决定拉伸效果的因素主要为图像的对比度及其亮度。利用灰度共生矩阵(gray-level co-occurrence matrix, GLCM)^[6]可以用来分析图像的灰度对比度。利用HSV空间的V分量就可以衡量其亮度。这里主要介绍对比度的衡量。GLCM是利用一个称为偏移模板的向量来对灰度图像进行运算的。向量表示对每个像素的计算范围,文中利用向量(0,1),表示计算当前像素的右邻像素。此向量对竖向灰度比较敏感,对应于车牌的竖向特征明显的特点。

计算过程为:首先将灰度图像分成若干灰度等级,每个像素的灰度值都对应一个灰度等级。设灰度分为 N 级,则GLCM矩阵的大小为 $N \times N$,元素 GLCM_{ij} 代表灰度等级为 i 其右邻像素灰度等级为 j 的所有像素

的个数总和。文中灰度等级设为 8, 则 GLCM 为一个 8×8 的矩阵。GLCM 的生成如图 1 所示。

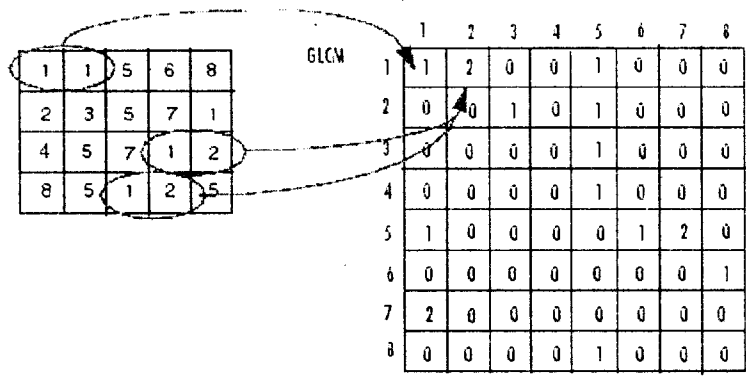


图 1 GLCM 的生成

在 GLCM 矩阵的基础上执行如下运算: 将 GLCM 归一为元素总和为 1 的矩阵, 每个元素 $P(i, j)$ 表示一个联合分布概率, 即从灰度等级为 i 的像素点离开某个固定位置(模版向量决定的)的点上的灰度等级为 j 的概率。对比度计算按照公式(4)进行。数字越大表明对比度越大。

$$\text{contrast} = \sum_{i,j} |i - j|^2 p(i, j) \quad (4)$$

对上面的实验进行了分类划分, 用不同拉伸参数对这 100 幅图像进行了一个全划分。发现只要分别利用 8 个拉伸参数就可以正确定位所有图像, 划分如表 2 所示。

通过表 2 得出结论: 每个拉伸参数对应一个图片对比度范围, 只要分析原灰度图像 O 的对比度并参考亮度, 采用相应的拉伸参数就可以对一幅图像进行很好的预处理, 实现方式如下: 将 O 划分成相同大小的子图, 分别求每个子图的亮度和对比度, 然后对比上表设置相应参数进行拉伸。最后再将图像合并起来生成图像 P1。对 O 进行子图分割时, 子图不能太大也不能太小。其大小最好能大于车牌字符但要小于整个车牌。实验证明 70×70 的子图比较合适。最后再对 P1 用灰度形态开运算的方法去除背景噪音, 减少背景噪音在二值化时对阈值的影响。实验发现利用饼状模板进行形态开运算的效果比较理想。而后用 Ostu 二值化图像, 它是一种求全局最佳阈值的方法。这时得到图像 P, 效果如图 2 所示。

表 2 不同对比度和亮度所对应的拉伸参数

对比度 \ 亮度	0~0.08	0.08~0.1	0.1~0.2	0.2~0.4	0.4~0.8	0.8~1.5	1.5~2.0	2.0 以上
大于 50	0.8	1.0	1.2	1.4	1.6	1.8	2	2.5
小于 50	0.8	1.0	0.8	1.0	1.4	1.8	2	2.5

2.2 BP 神经网络粗定位

算法采用 3 层网络, 输入层具有 200 个神经元, 输入值是每个像素值为 1 的 8-连接对象所占二值图像区域的 0、1 值, 8-连接对象需要首先归一化为 20×10 , 然后输入网络。挑选网络输入时可以先分析每个对象的长宽比例, 将不符合的对象过滤掉, 文中的高宽比例设置为 2 左右和 10 左右。用 tansig 作为隐层激活函数, 隐层只有一层, 隐含层经试验采用 10 个神经元已经足够。用 logsig 函数作为输出层激活函数, 输出层有一个神经元, 只能输出 1 和 0, 1 高表示车牌字符, 0 表示噪音。训练采用弹性 BP 算法。一般来说, 弹性 BP 算法的收敛速度要比梯度法更快。网络结构及应用方式如图 3 所示。

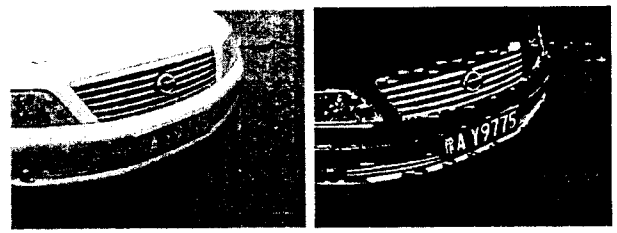


图 2 二值化效果图

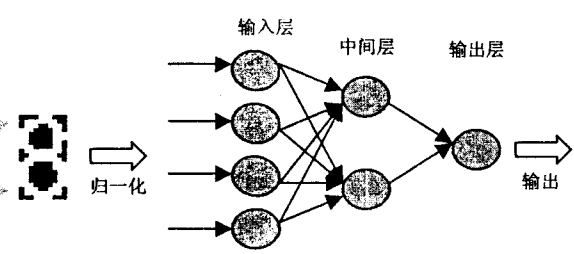


图 3 神经网络结构及应用图

算法只对车牌的数字和字母进行训练, 不需要进行汉字的训练。对每个数字和英文字母采用三个以上的样本进行训练, 样本是对字符从不同角度进行拍摄得到的。这样就让网络具有了认知车牌字符的能力, 网络还需要一定的噪音训练来排除背景对象, 实验发现, 照片中树木的枝叶、车灯以及竖向分布均匀的对象容易被网络识别为车牌字符, 所以算法着重对这些对象进行了提取用于网络训练。笔者通过实验提取了 90 幅车牌照片中的 1000 个字符以及噪音样本进行训练, 误差设为 $1e-3$ 时需要的训练次数在 2000 以内。

网络对二值图像进行分析后, 设定一个输出阈值 0.9, 大于这个值得则被认为很可能为车牌字符, 否则不予考虑。将这些字符记录到一个数组 C 中。C 的元素数量 n 在 10 到 100 之间。这样对图像分析后可以剔除大部分的背景噪音, 留下的输出为高的对象就是候选

车牌字符数组 C ,只需要对其中的每个元素进行分析。时间复杂度为 $O(n)$ 。

2.3 利用综合特征精确定位

对网络筛选出来的字符数组 C 再次进行分析,主要有三个标准:

I) HSV 图像空间分析,字符区域的 H 值经分析应该与蓝色、白色等车牌底色的分析相一致。

II) 车牌区域内应该至少有三个车牌字符能够被网络识别出来,也就是说数组 C 中的元素 C_i 一定范围内满足网络输出为高的对象应该有三个以上,这些对象肯定在 C 中,记为数组 P 。

III) 在 P 中,每个元素的高组成一个向量,其方差应该在很小的一个范围内。文中经试验设置为 $\varepsilon = 40$ 。下面分析每个标准的合理性:

I. 以蓝色车牌为例,底色为蓝色车牌的 H 分量在 195° 到 240° 之间,通过对 RGB 进行 HSV 转换后,利用聚类分析,聚类中心在 240° 的则满足标准。在 Matlab 中蓝色 H 值为 0.6,对每个 C_i 对应的区域进行 H 值的聚类分析。分为两类:背景蓝色的中心值大概在 0.6 (记为: H_{green}),另一类中心值与蓝色中心值相加(记为: H_{average}) 应该在 1.05 左右。这样每个区域就有两个值来表示: H_{green} 和 H_{average} 。利用式(5)来衡量每个区域与标准车牌区域的差值 dis 。满足 dis 小于 1 的区域为车牌字符。

$$\text{dis} = (H_{\text{green}} - 0.6)^2 + (H_{\text{average}} - 1.05)^2 \quad (5)$$

II. 本标准基于车牌字符分布特征,每个车牌经二值化后,最少会有三个字符完整地保留下来,被网络识别出。划定一个区域包含 C_i ,这个区域记为: χ ,如果 χ 里面包含三个以上网络输出值为高的 C 中的元素,则 C_i 满足第二个标准。这些元素被记录到一个数组 P 中。区域的划分有一个准则:尽量得小,但要包含完整的车牌。所以以 C_i 的大小作为标准,向左向右分别扩大 8 或者 40 倍(40 倍对应于字符 I) 的像素,向上向下分别扩大一倍的像素。

III. 由于车牌字符的大小是非常一致的,尤其是高都是一样的数值,所以第三个标准是合理的。但由于角度的变化和图像经处理后会有两个字符可能与车牌上方的白点发生粘连,并且有的字符的高会变小。所以其方差一般应该不为零,但是小于一定的数值 ε 。这里将 ε 设为 40。

经过上面三个标准的过滤,90% 以上的噪音会被过滤掉,留下的 C 中的元素就是车牌字符, P 中的元素即为车牌字符。这样区域 χ 就是车牌区域,其大小一般为车牌的两倍大。

3 实验结果

图 4 所示为车牌定位效果示例。数字图片表示经分析符合所有规则的 C_i 对应的 RGB 区域,第二列的图片表示区域 χ ,里面肯定包含车牌。到此车牌的精确定位已经完成。算法的处理时间主要耗费在图像的预处理上,试验硬件为奔腾 4 处理器,频率为 2.2GHz,如果图片的分辨率在 400×320 以下,算法执行速度为 1 秒左右,成功率可以达到 96%。

值得一提的是,经过文中所述的算法处理后,已经可以找到三个以上的车牌字符,使得后续分割算法的设计变得相对容易。



图 4 车牌定位示例图

参考文献:

- [1] 王宸昊,黄辉先,吴 翼,等.彩色汽车牌照的定位方法[J].测控技术,2006,25(6):69-71.
- [2] 陆 锋.基于改进的 BP 神经网络进行车牌定位的研究[J].苏州大学学报:工科版,2004,24(6):5-8.
- [3] Parisi R, Claudio E D D, Lucarelli G, et al. Car plate recognition by neural networks and image processing[C]//In Proc, IEEE, ISCAS. USA: [s. n.], 1998: 195-198.
- [4] 张晓清,王国文,曹海云.基于细化的手写汉字的笔段提取方法[J].哈尔滨工业大学学报,1999,31(5):107-110.
- [5] 丛 爽.面向 MATLAB 工具箱的神经网络理论与应用[M].合肥:中国科学技术大学出版社,2003.
- [6] Haralick R M. Textural features for image classification[J]. IEEE, 1973, 3(6): 610-621.

《计算机技术与发展》欢迎投稿,欢迎订阅。